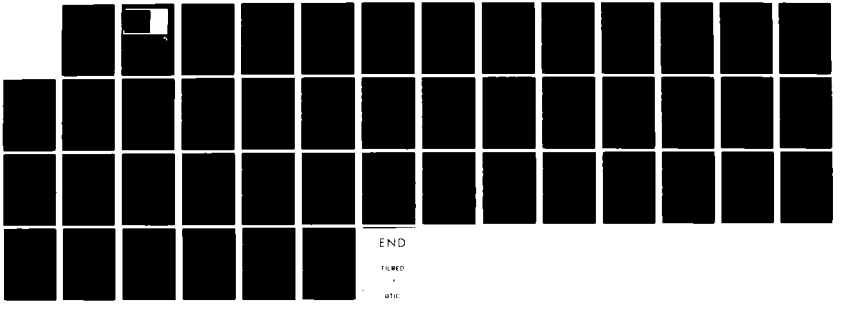
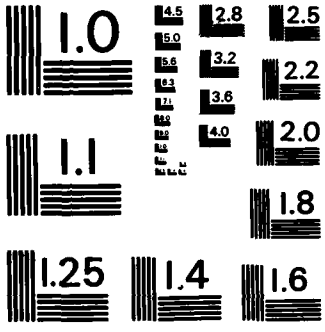
AD-A120 962 BOUNDARY CONDITIONS FOR SUPPRESSING RAPIDLY MOVING COMPONENTS IN HYPERBOL..(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER R L HIGDON JUL 82 MRC-TSR-2403 DAAG29-80-C-0041 1/1

UNCLASSIFIED F/G 20/4 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MRC Technical Summary Report #2403

BOUNDARY CONDITIONS FOR SUPPRESSING RAPIDLY MOVING COMPONENTS IN HYPERBOLIC SYSTEMS, I

Robert L. Higdon

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

July 1982

(Received May 5, 1982)

Approved for public release
Distribution unlimited

Sponsored by

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D.C. 20550

Office of Naval Research
Arlington, VA 22217

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

BOUNDARY CONDITIONS FOR SUPPRESSING RAPIDLY MOVING COMPONENTS IN HYPERBOLIC SYSTEMS, I

Robert L. Higdon

Technical Summary Report #2403
July 1982

ABSTRACT

This work is concerned with hyperbolic systems of partial differential equations for which certain of the associated propagation speeds are a great deal larger than the other propagation speeds. Our goal is to find boundary conditions which prevent rapidly moving waves from entering the given spatial domain. Conditions of this type are desirable in certain numerical computations arising in meteorology.

In order to find these conditions, we first transform the given system to an approximate diagonal form in such a way that each of the new dependent variables can be identified as a slow, incoming fast, or outgoing fast component of the solution. We then find local boundary conditions which suppress the incoming fast part. Pseudo-differential operators are used throughout the entire process. We consider only linear systems. In Part II of this work [7] these methods are applied in detail to the linearized shallow water equations.

The results of numerical tests of various boundary conditions are included in both papers. We also outline a method by which the conditions can be justified analytically.

AMS(MOS) Subject Classifications: 35L50, 65M99

Key Words: hyperbolic systems, absorbing boundary conditions, multiple time scales, reflection of singularities.

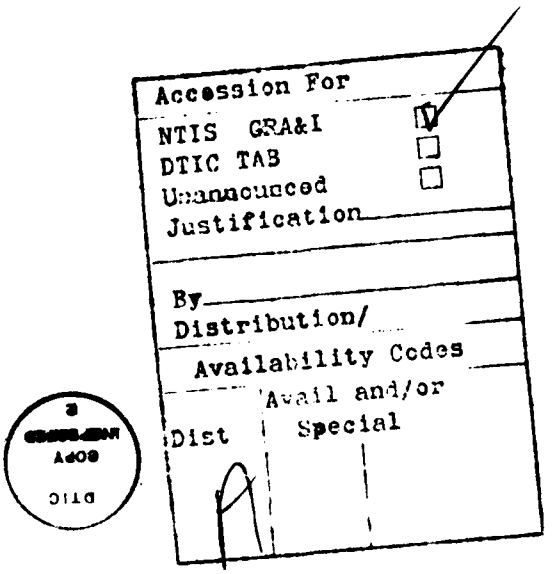
Work Unit Number 3 - Numerical Analysis and Computer Science

Research performed at Stanford University with partial support from Office of Naval Research Contract N00014-75-C-1132. Written at the Mathematics Research Center, University of Wisconsin - sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This materical is based upon work supported by the National Science Foundation under Grant No. MCS-7927062, Mod. 1.

SIGNIFICANCE AND EXPLANATION

When one computes a numerical approximation to the solution of a partial differential equation, it is sometimes necessary to restrict the computation to a domain which is only a portion of the domain on which the problem is naturally defined. This is done when one is interested in the behavior of the solution only on that part and when a computation over the entire domain would be prohibitively expensive.

In such cases a portion of the boundary of the computational domain represents merely the edge of the computation and corresponds to nothing physical. The task of finding suitable boundary conditions to impose on this part of the boundary can present some analytical and numerical difficulties. This work is concerned with one such difficulty which arises in limited-area meteorological computations. It is also related to the problem of finding conditions which prevent the nonphysical reflection of waves at an artificial boundary.



The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

BOUNDARY CONDITIONS FOR SUPPRESSING RAPIDLY MOVING COMPONENTS IN HYPERBOLIC SYSTEMS, I

Robert L. Higdon

1. INTRODUCTION

Hyperbolic partial differential equations are characterized by the fact that they propagate information at finite speed. For first order hyperbolic systems there may be several such propagation speeds, each corresponding to an eigenvalue of the principal symbol of the system. In the present work we consider systems for which the various speeds can have substantially different magnitudes. Systems of this type are sometimes said to have "multiple time scales".

Examples of such systems arise in the study of fluid dynamics. The shallow water equations and the Euler equations of gas dynamics admit certain modes associated with the movement of the fluid and certain other modes associated with the propagation of gravity waves and sound waves, respectively. If these waves move at speeds which are considerably greater than rate of flow of the fluid, then these systems have two time scales.

In this paper we consider initial-boundary value problems for such systems. The goal is to find boundary conditions which prevent rapidly moving high-frequency waves from entering the given spatial domain. That is, we will try to identify the portion of the solution which is entering the region at the fast speed, and we will then attempt to set this part of the solution equal to zero at the boundary. The purpose of these conditions is to prevent excessive errors in the boundary data from propagating rapidly into the

---

Research performed at Stanford University with partial support from Office of Naval Research Contract N00014-75-C-1132. Written at the Mathematics Research Center, University of Wisconsin - sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This material is based upon work supported by the National Science Foundation under Grant No. MCS-7927062, Mod. 1.

interior during numerical computations. In Section 2 of this paper we describe a physical problem in which boundary conditions of this type would be useful, and at the end of that section we mention a connection between this work and the construction of nonreflecting boundary conditions for hyperbolic systems.

In order to find the conditions desired here, we first transform the given system to an approximate diagonal form so that at high frequencies each of the new dependent variables can be identified as a slow, incoming fast, or outgoing fast component of the solution. Certain portions of this uncoupling process are similar to some ideas outlined by Engquist and Majda [1] for constructing absorbing boundary conditions for first order systems. After the system is uncoupled we find local boundary conditions which suppress the incoming fast part. Pseudo-differential operators are used extensively throughout the entire process. The discussion is limited to linear systems.

In Sections 3 and 5 of this paper we consider problems in one and several space dimensions, respectively. In order to simplify the discussions in these sections, we avoid using pseudo-differential operators and instead simulate their use with formal manipulations of Fourier transforms. The calculations with Fourier transforms give results which are valid for systems having constant coefficients, aside from a technical difficulty about the use of the transforms which is mentioned in Section 3.2. Pseudo-differential operators are used to settle this point and to generalize the methods of Sections 3 and 5 to systems which have variable coefficients. The generalization is outlined briefly in Section 4.1 and is developed in detail in Part II of the this work [7]. There we assemble the ideas of Section 5 and give a detailed development of boundary conditions for the linearized shallow water equations, or, equivalently, the two-dimensional Euler equations of gas dynamics.

The effects of various boundary conditions derived here can be analyzed using techniques from the theory of propagation of singularities for linear partial differential equations. This analysis is outlined in Section 4.2 of this paper, and it is discussed in greater detail in the thesis [6] on which these two papers are based. We also present the results of some numerical tests of the boundary conditions, in Section 3.4 for a model problem in one space dimension and in Part II for the shallow water equations.

2. PHYSICAL AND NUMERICAL ORIGINS OF THE PROBLEM

This work is motivated primarily by a situation which can arise when one computes numerical approximations to the solutions of hyperbolic systems which model the behavior of the earth's atmosphere. Under certain circumstances the boundary data available for these computations are substantially less accurate than the initial data which are available. The reason for this will be discussed below. This situation is unfortunate, since the large errors in the boundary data can generate comparable errors in the interior and thereby waste the extra accuracy contained in the initial data. The goal of the present work is to reduce as much as possible the extent to which this contamination takes place.

We propose to do this by imposing boundary conditions which prevent rapidly moving waves from propagating into the interior. The fast waves allowed by the systems in question are sound waves or gravity waves, and in the physical problem these waves represent only a small portion of the exact solution. But the fast modes are still able to carry substantial errors into the spatial domain, since the errors in the boundary data are not related to the exact solution and in general can excite all of the modes in the system. We would therefore like to identify the portion of the solution which is entering the domain at the fast speed and then set this portion equal to zero at the boundary. This would have the desired numerical effects and would also be physically realistic.

We will not attempt to prevent the propagation of error at the slow speed, since it is not realistic to assume that the slow part of the solution is equal to zero. However, it is still worthwhile to stop the rapidly moving portion, since in meteorological problems the fast and slow speeds can easily differ by a factor of five or ten to one.

The large errors in the boundary data are commonly found in limited area computations which are used to predict local atmospheric phenomena. In meteorological calculations a natural domain of definition for initial-boundary value problems is the entire earth's atmosphere. However, the size of this domain and the limitations of present-day computing machines require that global computations be done on meshes which are quite coarse. In current practice the grid spacing for such calculations is often one interval per two and a half degrees latitude and longitude. These calculations can give useful information about global phenomena, but the grid spacing is too coarse for predicting local phenomena.

It is common practice to perform additional computations over smaller regions with finer meshes so that these local phenomena can be resolved. For such a computation the spatial region is a cylinder in the atmosphere which is bounded by the earth's surface, the top of the atmosphere, and an artificial computational boundary. The artificial boundary merely defines the edge of the computation and represents nothing physical. The problem considered here arises when one tries to find suitable conditions to impose on this additional portion of the boundary. On this section it is necessary to prescribe values for variables which in some sense represent flow into the region, and for this one needs values of the solution on the boundary for times occurring after the initial time. This information is interpolated from the results of the global computation. Unfortunately, these results contain large errors resulting from the use of the coarse mesh, and they may also be affected by inadequate initial data on portions of the atmosphere. On the other hand, the initial data on the computational domain can be considered fairly reliable, since one would perform local computations only over a populated region where there is a dense network of observation stations capable of making accurate measurements.

This explains the substantial difference between the quality of the initial data and the quality of the boundary data.

The fast modes allowed by the system can cause another numerical difficulty besides the one discussed above. Because of the Courant-Friedrichs-Lewy condition, the fast modes can impose a severe restriction on the permissible time step for stable explicit difference approximations to the differential equation. In general, this presents the choice of either using an implicit difference method or an explicit method with very short time steps. Both choices are undesirable because of the computational expense which is involved. However, the physical insignificance of the fast waves makes it possible to use a certain reduced system which can be solved much more efficiently than the original system. A portion of the reduced system is elliptic and must be solved globally, but the remainder is hyperbolic with only slow propagation speeds. The use of such a system is based on a theory developed by Kreiss [9] for problems having multiple time scales, and it requires that the fast scale be properly filtered from the initial data. The implementation of this process is discussed by Gustafsson and Kreiss [5] and Gustafsson [3],[4].

This method does not quite solve the problem discussed in this paper, since the elliptic part of the reduced system can cause certain portions of the errors in the boundary data to propagate instantaneously into the interior. In this paper we will be concerned mainly with the high-frequency portion of the solution, since numerical errors are mainly of this type. Errors having high frequencies in the tangential variables will be damped by the elliptic part of the system, but this is not the case for errors in t. The methods of this paper are oriented mainly toward high frequencies in time.

It might seem that we could best deal with the problems described in this section by modifying the system of differential equations so as to prohibit solutions containing rapidly moving waves. However, it is not known how to do this and still retain a system which is mathematically well-behaved and which is a sufficiently accurate model of the atmosphere to be useful for meteorological computations.

We conclude by noting that the work in this paper is relevant to the problem of finding boundary conditions which prevent the nonphysical reflection of waves by an artificial boundary. A wave leaving the spatial domain will be reflected back into the interior if and only if the boundary conditions allow the outgoing modes to stimulate the incoming modes. The linkage between the modes is broken if the system is uncoupled and the incoming modes are set equal to zero at the boundary. Conditions of this type are realistic if the solution consists entirely of outgoing waves.

3. SYSTEMS IN ONE SPACE DIMENSION

3.1 General Remarks

We will consider the hyperbolic system

$$\frac{\partial}{\partial t}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a & \\ & b \end{pmatrix}\frac{\partial}{\partial x}\begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} \tag{3.1}$$

for $0 < x < 1$, $t > 0$. This can also be written $w_t = Aw_x + Cw$, where $w = (u,v)^T \in R^2$. The entries in A and C are functions of x and t.

In order to simplify the notation we have chosen a system having two scalar components. It can be seen easily that the ideas presented in this section work equally well for systems having several components.

There is no loss of generality in assuming that A is diagonal. The system is hyperbolic, so A has real eigenvalues and a complete set of eigenvectors. If A is not diagonal, then a suitable similarity transformation and change of dependent variables can be made to bring the system to diagonal form.

We assume $|a| << |b|$ and $a < 0$, $b < 0$. The first assumption guarantees the presence of propagation speeds having substantially different magnitudes. The second assumption is made for the sake of definiteness. It also contains the assumption that $\det A \neq 0$, i.e., that the boundary is noncharacteristic.

The problem is to identify the "fast" part of the solution of the system and then find boundary conditions which suppress this as much as possible. To some degree this can be done via the method of characteristics. If C were diagonal, then the system (3.1) would consist of

two independent equations for u and v. These are ordinary differential equations along the characteristics curves $\frac{dx}{dt} = -a$ and $\frac{dx}{dt} = -b$, respectively. Initial values for these equations are provided by initial data ($t = 0$) and boundary data ($x = 0$, in this case) for (3.1). Since $|b| >> |a|$, v is the component which propagates information at the fast speed, and it would therefore suffice to set $v = 0$ at the boundary.

However, if C is not diagonal, then this boundary condition may not be adequate. In this case u can act as a forcing function in the equation for v, and in general u is nonzero. The boundary data could thus influence the solution in the interior at the fast speed by first influencing u, which in turn would force v.

We would therefore like to have boundary conditions which are more effective than the one mentioned above. These conditions can be obtained by transforming the system so as reduce the coupling found in the lower order term. This can be thought of as a process of identifying more precisely the quantity which moves slowly and the quantity which moves rapidly. Refined boundary conditions can then be derived by setting the new fast variable equal to zero whenever permissible and then expressing this condition in terms of the original unknowns u and v. In the next section we discuss a method for reducing the coupling in the system, and in Section 3.3 we produce boundary conditions from the results of the uncoupling.

3.2 A Method for Reducing the Coupling in Lower-Order Terms

This method was used by Taylor [12] to reduce the coupling in systems of pseudo-differential equations in which the leading symbol is already known to be in block-diagonal form. It is essentially a simple

perturbation argument, valid for large frequencies, which is disguised by the language of pseudo-differential operators. It is similar to some uncoupling methods used by Kreiss [8] and O'Malley and Anderson [11]. In this sub-section we outline the technique using Fourier transforms in a formal manner. A general version of the perturbation method is given in the Appendix.

Before this method can be used, the leading order part of the system must be brought to diagonal form, or at least to a suitable block-diagonal form. For systems in one space dimension this step is trivial, and it has already been accomplished for the system (3.1). In several dimensions this step is rather complicated. It will be discussed at length in Section 5.

The system considered here is $w_t = Aw_x + Cw$. We will first apply a Fourier transform in order to reduce it to a system of ordinary differential equations, and we will then apply the perturbation argument. In this section we assume that the system has constant coefficients.

We will not apply a transform in x because this would require information about the solution away from the boundary. That would not be appropriate in a discussion of boundary conditions. Instead, we will use a Fourier transform in t. The use of such a transform is complicated by the fact that the solution can grow too rapidly as $t \to \infty$ for the transform to be defined in the usual sense, but it is possible to remedy this problem with a localization argument which uses properties of pseudo-differential operators. Details are given in [6]. The growth problem could also be solved by using a Laplace transform, but we will not do this because the time derivative would cause the initial data to be introduced into the transformed equations. That would be an undesirable

complication. It is convenient to use pseudo-differential operators to handle this problem since we were going to introduce these operators, anyway, in order to treat systems having variable coefficients.

The system (3.1) can be written as

$$w_x = A^{-1}w_t - A^{-1}Cw \tag{3.2}$$

In terms of components this is

$$\frac{\partial}{\partial x}\begin{pmatrix}u\\v\end{pmatrix} = \begin{pmatrix}a^{-1} & \\ & b^{-1}\end{pmatrix}\frac{\partial}{\partial t}\begin{pmatrix}u\\v\end{pmatrix} - \begin{pmatrix}a^{-1}c_{11} & a^{-1}c_{12}\\ b^{-1}c_{21} & b^{-1}c_{22}\end{pmatrix}\begin{pmatrix}u\\v\end{pmatrix}. \tag{3.3}$$

Introduce formally the Fourier transform in t. Let ξ be the dual variable, and let $\hat{u},\hat{v},\hat{w}$ be the transforms of u,v,w. If the coefficients in the equation are constant, (3.2) becomes

$$\begin{aligned}\hat{w}_x(x,\xi) &= i\xi A^{-1}\hat{w} - A^{-1}C\hat{w}\\ &= i\xi(A^{-1} - \frac{1}{i\xi}A^{-1}C)\hat{w}\\ &= i\xi R(i\xi)\hat{w}.\end{aligned} \tag{3.4}$$

When ξ is large the matrix $R(i\xi)$ is a perturbation of the diagonal matrix A^{-1}. We will find a similarity transformation which reduces the coupling caused by the off-diagonal elements, and we will then use this transformation to make a change of dependent variable.

Let $Q(i\xi) = I + (i\xi)^{-1}M$, where M is a matrix to be determined. For large ξ, Q^{-1} exists and has the expansion

$$Q^{-1} = I - \frac{1}{i\xi}M + O(\frac{1}{\xi^2}).$$

Using (3.4), we have

$$QRQ^{-1} = (I + \frac{1}{i\xi}M)(A^{-1} - \frac{1}{i\xi}A^{-1}C)(I - \frac{1}{i\xi}M + O(\xi^{-2}))$$

(3.5)

$$= A^{-1} + \frac{1}{i\xi}(MA^{-1} - A^{-1}M - A^{-1}C) + O(\xi^{-2}).$$

The coupling is of order -2 if $MA^{-1} - A^{-1}M - A^{-1}C$ is diagonal, i.e., if

$$(3.6)\quad \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}\begin{pmatrix} a^{-1} & \\ & b^{-1} \end{pmatrix} - \begin{pmatrix} a^{-1} & \\ & b^{-1} \end{pmatrix}\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} - \begin{pmatrix} a^{-1}c_{11} & a^{-1}c_{12} \\ b^{-1}c_{21} & b^{-1}c_{22} \end{pmatrix}$$

is a diagonal matrix. We therefore set to zero the off-diagonal entries in (3.6).

$$(3.7)\quad m_{12}b^{-1} - a^{-1}m_{12} - a^{-1}c_{12} = 0 \quad \text{(row 1, column 2)}$$

$$m_{21}a^{-1} - b^{-1}m_{21} - b^{-1}c_{21} = 0 \quad \text{(row 2, column 1)}$$

This yields

$$m_{12} = \frac{a^{-1}c_{12}}{b^{-1} - a^{-1}} = \frac{bc_{12}}{a-b}$$

$$m_{21} = \frac{b^{-1}c_{21}}{a^{-1} - b^{-1}} = \frac{ac_{21}}{b-a}\ .$$

No conditions are imposed on the values of m_{11} and m_{22}. For convenience we take these to be zero. The matrix Q is then given by $Q = I + (i\xi)^{-1}M$, or

$$(3.8)\quad Q = I + \frac{1}{i\xi}\begin{pmatrix} 0 & \frac{bc_{12}}{a-b} \\ \frac{ac_{21}}{b-a} & 0 \end{pmatrix}$$

Equation (3.5) becomes

$$
\begin{aligned}
\Omega R \Omega^{-1} &= \Omega(A^{-1} - (i\xi)^{-1}A^{-1}C)\Omega^{-1} \\
&= A^{-1} + \frac{1}{i\xi}\begin{pmatrix} -a^{-1}c_{11} & 0 \\ 0 & -b^{-1}c_{22} \end{pmatrix} + O\left(\frac{1}{\xi^2}\right)
\end{aligned} \tag{3.9}
$$

We now use the similarity transformation to reduce the coupling in equation (3.4). This system can be written in the form

$$
\frac{\partial}{\partial x}(\Omega\hat{w}) = \Omega(i\xi A^{-1} - A^{-1}C)\Omega^{-1}(\Omega\hat{w}) \tag{3.10}
$$

If we let $\hat{w}_1 = \Omega\hat{w}$ and use (3.9), the system (3.10) becomes

$$
\frac{\partial \hat{w}_1}{\partial x} = \left\{ i\xi \begin{pmatrix} a^{-1} & \\ & b^{-1} \end{pmatrix} - \begin{pmatrix} a^{-1}c_{11} & \\ & b^{-1}c_{22} \end{pmatrix} + O\left(\frac{1}{\xi}\right) \right\} \hat{w}_1 \tag{3.11}
$$

When ξ is large, the coupling caused by the lower order term in (3.11) is weaker than the coupling in the original system (3.4). This means that for large ξ we can identify more precisely the rapidly moving part of the solution and do a better job of suppressing it. This will be done in the next sub-section. The restriction to high frequencies is not a serious one, since the goal of this work is to suppress the effect of numerical errors, which are mainly high frequency phenomena.

This method can be applied repeatedly to reduce even further the coupling at high frequencies. To reduce the coupling to $O(\xi^{-2})$, we would multiply (3.11) by a matrix of the form $I + (i\xi)^{-2}M_2$, and then determine M_2 in the same way that we found the matrix M above. In general, to reduce the coupling from $O(\xi^{-n+1})$ to $O(\xi^{-n})$, we would use

a multiplier of the form $I + \xi^{-n}M_n$. The details of this process involve no new ideas and will not be given here.

The method has been presented for 2×2 matrices. It can also be used to reduce the coupling in square matrices of any order which are perturbations of block-diagonal matrices. In this case the equations analogous to (3.7) can be solved provided the diagonal blocks corresponding to a^{-1} and b^{-1} have disjoint spectra. This generalization will be discussed in the Appendix.

3.3 Boundary Conditions

We now use the results of the uncoupling to find boundary conditions which suppress the fast part of the solution. The new dependent variable is defined by $\hat{w}_1 = Q\hat{w}$. By (3.8) this is

$$\begin{pmatrix} \hat{u}_1(x,\xi) \\ \hat{v}_1(x,\xi) \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{i\xi}\left(\frac{bc_{12}}{a-b}\right) \\ \frac{1}{i\xi}\left(\frac{ac_{21}}{b-a}\right) & 1 \end{pmatrix} \begin{pmatrix} \hat{u}(x,\xi) \\ \hat{v}(x,\xi) \end{pmatrix} \tag{3.12}$$

v_1 is our new notion of what constitutes the rapidly moving part of the solution. For large ξ it is a perturbation of the fast characteristic variable v. To suppress the fast part we set $v_1 = 0$ at $x = 0$, i.e.,

$$\frac{1}{i\xi}\left(\frac{ac_{21}}{b-a}\right)\hat{u} + \hat{v} = 0 \quad \text{at} \quad x = 0. \tag{3.13}$$

To obtain local boundary conditions we multiply (3.13) by $i\xi$ and then apply an inverse Fourier transform. The result is

(3.14) $$\frac{\partial v}{\partial t} + \left(\frac{ac_{21}}{b - a}\right)u = 0 \quad \text{at} \quad x = 0.$$

When the coefficients in the system depend on t, this derivation is obviously not valid. However, it is possible to use pseudo-differential operators to show that the boundary condition (3.14) still has some desirable effects. This is done in [6], and in Part II [7] we do something analogous for the linearized shallow water equations.

In order to define a well-posed problem, we must also prescribe a boundary condition for the slow part of the solution of (3.1). In this case it is necesary to prescribe a value at $x = 0$. One possible condition is

$$u = \text{given function} \quad , \text{ for } x = 0.$$

We might also use the result (3.12) of the uncoupling. That is, we could prescribe a value for $\hat{u}_1$, clear denominators, and then obtain a boundary condition analogous to (3.14). However, there seems to be little value in doing this, so this possibility will be disregarded.

We therefore propose the boundary conditions

(3.15) $$\frac{\partial v}{\partial t} + \left(\frac{ac_{21}}{b - a}\right)u = 0$$
$$u = g, \quad \text{for } x = 0,$$

where g is a given function of t. An integration of the first equation shows that $v(0,t)$ is defined in terms of g and the initial value $v(0,0)$. The conditions (3.15) thus prescribe values for the characteristic variables u,v on the portion of the boundary where the characteristics enter the region. This implies that the initial-boundary value problem is well-posed.

3.4 Numerical Computations

In this sub-section we present the results of some numerical computations which indicate the value of reducing the coupling caused by lower order terms. We consider the system

$$\frac{\partial}{\partial t}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -1 & \\ & -5 \end{pmatrix}\frac{\partial}{\partial x}\begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 10 & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} \tag{3.16}$$

for $0 < x < 1$ and $t > 0$. This is the system (3.1), where $a = -1$, $b = -5$, $c_{21} = 10$, and the other c_{ij} are zero.

We compare the boundary conditions

$$\begin{aligned} v &= 0 \\ u &= g \end{aligned} \qquad (x = 0) \tag{3.17}$$

and

$$\frac{\partial v}{\partial t} + \left(\frac{ac_{21}}{b - a}\right) u = 0 \qquad (x = 0)\ . \tag{3.18}$$

$$u = g$$

Here g is a given function of t. The simpler conditions (3.17) are derived from the method of characteristics discussed in Section 3.1. The first condition in (3.18) was obtained from the results of the uncoupling process.

In our computation the system is approximated by the leap frog difference scheme. The function g in the boundary conditions is equal to a half period of a sine wave which is extended by zero. A forward difference is used to approximate the derivative in (3.18). The surfaces pictured in Figures 3.2 and 3.3 are graphs of $\sqrt{u^2 + v^2}$ as a function of x and t. In Figure 3.1 we illustrate the configuration of these surface plots.

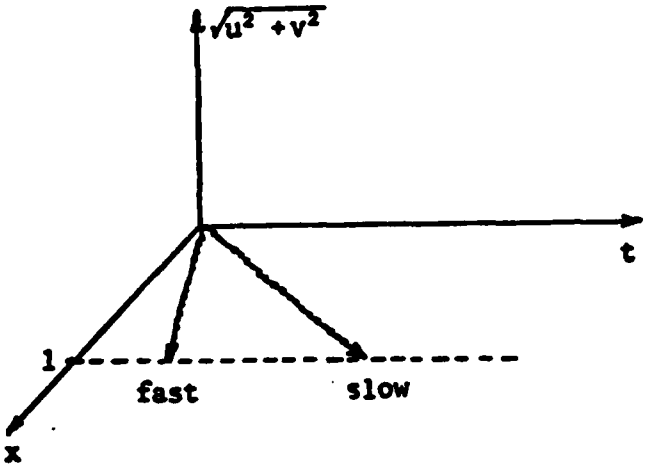


Figure 3.1

In the computations we set the solution equal to zero when $t = 0$. The nonzero part of the solution is thus due entirely to the nonzero boundary data, so it is possible to study the influence of the boundary data by examining the size of the solution in various parts of the (x,t) plane. The solution corresponding to the simple boundary condition (3.17) is graphed in Figure 3.2, and the solution corresponding to the more refined condition (3.18) is given in Figure 3.3. It is clear from the figures that the second condition is much more effective in suppressing the fast part of the solution.

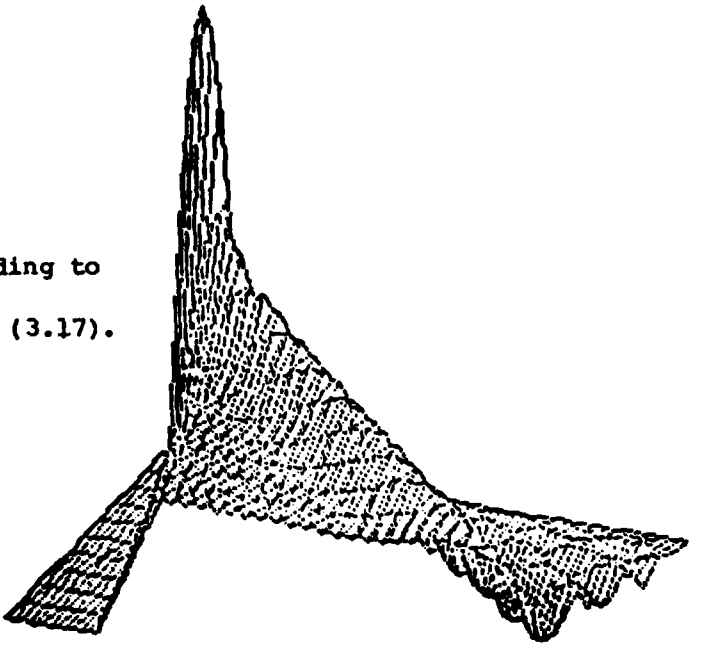


Figure 3.2. Solution corresponding to boundary condition (3.17).

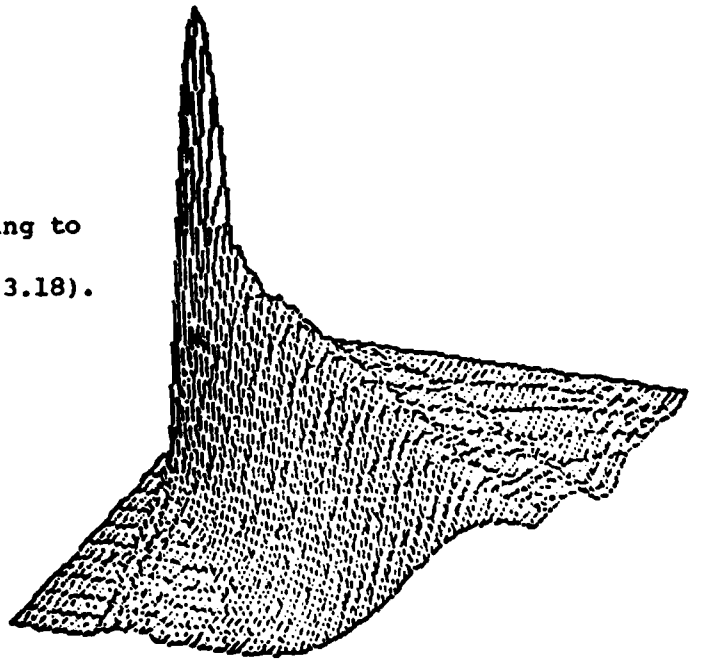


Figure 3.3. Solution corresponding to boundary condition (3.18).

4. VARIABLE COEFFICIENTS; ANALYSIS OF THE BOUNDARY CONDITIONS

In this section we discuss a couple of theoretical questions which are suggested by the calculations given in Section 3.

4.1 Generalization to Systems Having Variable Coefficients.

Problems with variable coefficients can be treated by using pseudo-differential operators in a manner which is analogous to the development in the preceding section. This is done for the one-dimensional problem in [6], and in Part II we use these operators on an example in two space dimensions. The purpose of the present subsection is to give a brief description of the ideas behind the generalization.

A pseudo-differential operator has the form

$$Pu(t) = \int e^{i\xi t} p(t,\xi)\, \hat{u}(\xi)\, d\xi \tag{4.1}$$

For example, if the symbol p is a polynomial in ξ with coefficients depending at most on t, then P is also a differential operator. In the Appendix in Part II we describe a class of permissible symbols and summarize some formal properties of these operators.

During the calculations in Section 3 we multiplied the Fourier transform by functions of the dual variable ξ. In effect, we were applying pseudo-differential operators whose symbols were independent of t. In this case there is no need to carry out the integration which inverts the Fourier transform, and the introduction of a name for these operators contributes nothing to the discussion.

When the coefficients in the system depend on t, the manipulations in Section 3 are obviously not valid. However, it is possible to carry

out analogous calculations which are valid in this case. Instead of multiplying the Fourier transform by functions of the dual variable, one applies operators defined in terms of the inverse Fourier transform, as in (4.1). These calculations involve the use of expansions which are valid asymptotically as $|\xi| \to \infty$, and, roughly speaking, the leading order terms in these expansions are the results which one would obtain by calculating formally with Fourier transforms as in Section 3. This is to be expected, since variable coefficients appear almost constant to a wave whose frequency is sufficiently high. The expressions for the lower order terms in the expansions involve derivatives of the coefficients. Pseudo-differential operators can thus be regarded as a formalism for studying the high-frequency asymptotic behavior of the solution, and in the present work they are used both to uncouple the system and to derive boundary conditions from the results of the uncoupling.

4.2 Estimate of the Size of the Fast Part of the Solution

Section 3.4 provides numerical evidence of the effectiveness of the boundary conditions which were derived in Sections 3.2 and 3.3. Here we outline a method for justifying the boundary conditions analytically. Details are given in [6]. This method also works for problems in several space dimensions.

When a system with constant coefficients is uncoupled as in Section 3, one can obtain an ordinary differential equation in x for the Fourier transform of the portion of the solution which is entering the region at the fast speed. There is an equation for each frequency. The solution can be estimated in a manner which is essentially the same as the one used to obtain energy estimates for hyperbolic equations. The result is a

bound for the incoming fast part in terms of its value at the boundary and the magnitude of the forcing caused by other components in the system. The choice of boundary conditions and the fact that the forcing is small imply that the incoming fast part of the solution must be small.

When the coefficients in the system are not constant, it is not possible to examine the solution one frequency at a time. However, one can still consider the restriction of the solution to a conical neighborhood of a bicharacteristic curve. This amounts to a study of the propagation along a ray of a family of neighboring high frequencies. Energy estimates for the restriction yield results which are analogous to those found in the case of constant coefficients. The construction of a suitable restriction operator is given in Nirenberg [10, p.44].

In several dimensions the restriction process enables one to examine the effects of the boundary conditions at various angles of incidence to the boundary. The utility of this is implied by the discussion appearing in the next section.

5. SYSTEMS IN SEVERAL SPACE DIMENSIONS

In this section we generalize the methods of Section 3 to systems in several dimensions. The main point of discussion is the problem of uncoupling the leading order part of the system. The lower order terms can be treated in the same manner as before. A summary of the uncoupling process is given in the last sub-section.

5.1 Assumptions

We consider the hyperbolic system

$$w_t = Aw_x + Bw_y + Cw \tag{5.1}$$

for $x > 0$, $y \in \mathbf{R}$. Here $w(x,y,t) \in \mathbf{R}^n$, and A,B, and C are real $n \times n$ matrices. We assume that A is nonsingular, i.e., that the boundary is noncharacteristic, and without loss of generality we also assume that A is diagonal. In order to simplify the notation we have chosen a system in two space dimensions. However, it will be clear that the discussion is equally valid for systems in higher dimensions where $x > 0$ and $y \in \mathbf{R}^k$ for $k \geq 2$.

There is no serious loss of generality in assuming that the spatial domain is a half-space. If the given domain does not have this form but still has a smooth boundary, then it is possible to localize the problem with a partition of unity and then map each boundary portion into the boundary of a half-space. In the new coordinates the problem will have the form given above.

The system (5.1) has been assumed to be hyperbolic. In this paper

this means that for every real ζ and ω, the symbol

$$\zeta A + \omega B \tag{5.2}$$

has real eigenvalues and a complete set of real eigenvectors. There will be no need to assume that A and B are symmetric or that the system is strictly hyperbolic.

In order to have a system with at least two time scales, we assume that certain eigenvalues of the symbol (5.2) are substantially greater in magnitude than the others. In the case of the linearized shallow water equations this symbol has eigenvalues $-u\cdot\sigma$ and $-u\cdot\sigma + c|\sigma|$, where $u = (u_1,u_2)$ is the velocity of the flow about which the system has been linearized, and σ is the vector of dual variables (ζ,ω). If $|u|<<c$, then this system has two time scales. There is a similar set of eigenvalues for the three-dimensional, five-component Euler system for gas dynamics. In this case the small eigenvalue has multiplicity three.

5.2 An Overview of the Problem

We wish to transform the system (5.1) to an approximate diagonal form, or at least block diagonal form, so that each of the new dependent variables can be identified as a slow, incoming fast, or outgoing fast portion of the solution. We will then attempt to set the incoming fast components equal to zero.

The immediate goal is to uncouple the leading order terms in the system (5.1). After this has been accomplished, we can use the methods of Section 3 to reduce the coupling caused by the lower order terms. It would suffice to obtain a block diagonal form for the system, instead of

diagonal form, since there is no need to separate various incoming fast components or various incoming or outgoing slow components. This situation could occur with the Euler equations, for example.

We have assumed that the matrix A in (5.1) is already in diagonal form. This involves no loss of generality, since if A is not in that form we can find a similarity transformation which makes it diagonal and then adopt a suitable change of dependent variable. Unfortunately, it is not true in general that this transformation can also uncouple the matrix B. It is therefore necessary to do something extra if we want to uncouple the entire principal part of (5.1).

In the case of constant coefficients it may be tempting to use Fourier transforms in x and y. This would yield the equation

$$\hat{w}(\zeta,\omega,t) = (i\zeta A + i\omega B)\hat{w} + C\hat{w}.$$

The leading symbol of this equation can be diagonalized easily because it is a scalar multiple of the symbol (5.2) discussed earlier. However, the use of Fourier transforms in x requires the use of information about the solution away from the boundary $x = 0$, and this is not appropriate in a discussion of boundary conditions. We will instead apply Fourier transforms in time and in the tangent variable y. The use of the transform in t can be justified in the manner indicated in Section 3.2.

The system (5.1) can be written in the form

$$w_x = A^{-1}w_t - A^{-1}Bw_y - A^{-1}Cw \tag{5.3}$$

Let $\hat{w}(x,\omega,\xi)$ denote the Fourier transform of w with respect to y and t for fixed x. Equation (5.3) implies

$$\hat{w}_x(x,\omega,\xi) = (i\xi A^{-1} - i\omega A^{-1}B)\hat{w} - A^{-1}C\hat{w}. \tag{5.4}$$

It is necessary to determine the values of ω and ξ for which the symbol

$$\xi A^{-1} - \omega A^{-1}B \tag{5.5}$$

can be diagonalized, or brought to block diagonal form, and we must determine whether such an uncoupling can produce a transformed system in which each component of the dependent variable can be identified as slow, incoming fast, or outgoing fast. The answers to these questions are not immediately obvious, since we are applying transforms in the nonstandard variables y and t.

5.3 Properties of the Symbol (5.5)

In this sub-section we consider the questions posed in the preceding paragraph, interpret the results physically, and discuss methods for obtaining explicit formulas for transformations which bring about approximate diagonal forms for (5.5).

We must first consider the eigenvalues and eigenvectors of (5.5). Suppose that ζ is a real eigenvalue and that v is a corresponding eigenvector. This means that

$$(\xi A^{-1} - \omega A^{-1}B)v = \zeta v. \tag{5.6}$$

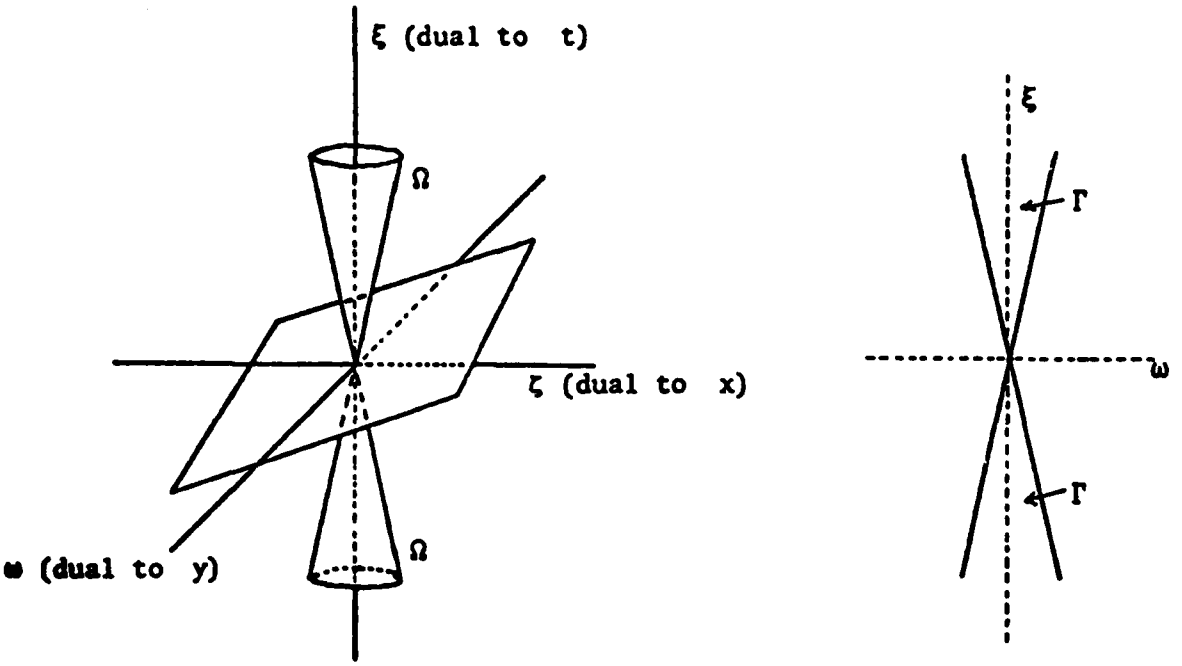
If we multiply by A and rearrange the terms, the result is

$$(\zeta A + \omega B)v = \xi v. \tag{5.7}$$

The matrix $\zeta A + \omega B$ is the symbol (5.2) which we would obtain by writing the system in the more common form (5.1) and then applying Fourier trans-

forms in the usual variables x and y. According to (5.6) and (5.7), this symbol imposes the same relations between the dual variables ζ, ω, and ξ as the symbol (5.5), and it is possible to find the eigenvectors of one symbol by examining the eigenvectors of the other. The difference between the two situations is that in one case the variable ζ is treated as a function of ω and ξ, and in the other case ξ is treated as a function of ζ and ω. This correspondence between the two symbols will be exploited in studying (5.5), since at this point in the discussion we know a great deal more about (5.2) than we do about (5.5).

In order to have a system with multiple time scales we have assumed that certain eigenvalues ξ of (5.2) are considerably larger than the others. An example of such a set of eigenvalues is graphed in Figure 5.1(a). In this example there are two relatively large eigenvalues and one smaller eigenvalue for each ζ and ω. This is the configuration for the shallow water equations, and it is similar to the configuration for the Euler equations of gas dynamics. In the latter case the small eigenvalue has multiplicity three. Throughout this discussion we will assume that the largest eigenvalues of (5.2) have graphs which are narrow cones, although not necessarily right circular cones. These eigenvalues must come in pairs, since if (ζ,ω,ξ) is a solution of (5.7) then so is $(-\zeta,-\omega,-\xi)$. These large eigenvalues correspond to rapidly moving waves. The fact that the graphs are not necessarily right circular cones means that the speed can vary somewhat with the direction of propagation. We will denote by Ω the double cone which corresponds to the largest eigenvalues of (5.2), and we will denote by Γ the projection of Ω onto the (ω,ξ) space. These are labeled in Figure 5.1.



(a) Graph of the relations (5.6), (5.7)

(b) Projection onto the (ω,ξ) space

Figure 5.1

We now use Figure 5.1(a) to study the behavior of the eigenvalues ζ of the symbol (5.5), $\xi A^{-1} - \omega A^{-1}B$. First of all, it is apparent that the number of real eigenvalues must vary with the position of (ω,ξ). If (ω,ξ) lies in Γ, then there are two real values of ζ which are associated with the surface Ω. As (ω,ξ) approaches the boundary of Γ, these values of ζ coalesce, and when (ω,ξ) leaves Γ the eigenvalues leave the real axis and form a pair of complex conjugates. The eigenvalues cannot be real, since for any real ζ the point (ζ,ω,ξ) must lie on one of the surfaces in Figure 5.1(a). They are complex conjugates because they are eigenvalues of a real matrix.

It is safe to assume that for (ω,ξ) in a neighborhood of Γ there is no problem in solving for the values of ζ associated with surfaces

different from Ω. In the case of the shallow water equations the other ζ satisfies the equation $\xi = -u_1\zeta - u_2\omega$. We have assumed that the matrix A in (5.1) is nonsingular, which in this case is equivalent to saying $u_1 \neq 0$. It is therefore possible to solve for ζ in terms of ω and ξ, whether or not (ω,ξ) is in Γ. A similar situation holds for the Euler equations of gas dynamics.

We now characterize the behavior of (5.5) when (ω,ξ) lies in Γ. It will become clear a little later that this is the only portion of the (ω,ξ) space in which we are really interested.

Proposition. If (ω,ξ) is in Γ, then the symbol (5.5), $\xi A^{-1} - \omega A^{-1}B$, has real eigenvalues and a complete set of real eigenvectors. This is not the case if (ω,ξ) is not in Γ. The eigenvectors can be determined from those of the symbol (5.2), $\zeta A + \omega B$.

Proof. Equation (5.6) and (5.7) show that the eigenvectors of (5.2) are also eigenvectors of (5.5). We know that (5.2) has a complete set of real eigenvectors corresponding to fixed (ζ,ω) and various eigenvalues ξ. We want to show the same thing for (5.5), for fixed (ω,ξ) in Γ, and various eigenvalues ζ.

Suppose that (ω,ξ) in Γ, and let $\zeta_1,\ldots,\zeta_m$ denote the eigenvalues of (5.5). For each ζ_j choose a basis B_j for the eigenspace of $\zeta_j A + \omega B$ corresponding to eigenvalue ξ. We are allowing for the possibility that the symbol (5.2) might have mutliple eigenvalues. The elements of B_j are also eigenvectors of (5.5) corresponding to the eigenvalue ζ_j. We claim that the union of the B_j is a complete set of vectors. There are clearly enough of these

vectors. The fact that they are linearly independent follows from an argument which is essentially the one which shows that eigenvectors corresponding to distinct eigenvalues are linearly independent. This completes the proof. ∎

The matters discussed in this section can be given a physical interpretation. Suppose that the coefficients in (5.1) are constant, and let $C = 0$. This gives the system

$$w_t = Aw_x + Bw_y. \tag{5.8}$$

If we insert a plane wave solution $v \exp(i\zeta x + i\omega y + i\xi t)$ into (5.8), where v is a vector, the result is $\xi v = (\zeta A + \omega B)v$. This is the condition (5.7) which was discussed earlier. The surfaces in graphs like Figure 5.1(a) therefore define the set of all possible frequencies for plane wave solutions to (5.8). It is apparent that the rapidly moving waves are associated with Ω, which is why we are interested in the behavior of (5.5) only for (ω,ξ) in Γ. This analysis is also valid when $C \neq 0$, provided we are considering high-frequency waves.

In graphs like Figure 5.1(a) there is a particular wave speed associated with each surface which defines ζ as a function of ω and ξ. This implies that it is possible to separate fast waves from slow waves by diagonalizing the symbol (5.5). It is also necessary to detect the directions in which the fast waves are moving, since we want the diagonal form to distinguish incoming and outgoing fast waves. This separation is possible. The vector group velocity for plane wave solutions of (5.8) is equal to the gradient of $-\xi$ with respect to (ζ,ω). From this it follows that each determination of ζ on each section of Γ can be identified with motion into or out of the domain

$x > 0$, and it is also clear that the edge of Γ corresponds to tangentially moving waves. By properly defining the branches of ζ on the two sections of Γ, we can therefore separate the fast part of the solution into incoming and outgoing components. This justifies our decision to seek diagonal form for the symbol (5.5).

We need to say a little more about the directions in which the various waves propagate. When we seek explicit formulas for uncoupling the symbol (5.5), we will introduce approximations which are valid asymptotically as $\frac{\omega}{\xi} \to 0$. The case $\frac{\omega}{\xi} = 0$ corresponds to normal phase velocity for a wave of the form $\exp(i\zeta x + i\omega y + i\xi t)$, and it corresponds approximately to a normal group velocity for fast waves. If Ω were a right circular cone, then these velocities would coincide. The approximations will thus lead to boundary conditions which work well for fast waves traveling in directions having sizeable normal components, but they may not work well for waves moving in directions which are nearly tangential, i.e., for (ω,ξ) near the edge of Γ. These tangential waves do not present any real problem, since they cannot influence the interior very rapidly. The approximation schemes are therefore worth using.

One such scheme is the perturbation method given in the Appendix. The symbol (5.5) can be written in the form $\xi(A^{-1} - \frac{\omega}{\xi}A^{-1}B)$. When $|\frac{\omega}{\xi}|$ is small, the matrix $A^{-1} - \frac{\omega}{\xi}A^{-1}B$ is a perturbation of the diagonal matrix A^{-1}, and the perturbation method can be used to uncouple it to higher powers in $\frac{\omega}{\xi}$. From the discussion in the Appendix it is apparent that in the case of multiple eigenvalues this method cannot guarantee diagonal form, but instead can give a suitable block diagonal form.

Another way to uncouple (5.5) is to compute its eigenvectors directly. One could either work directly with (5.5) or instead find

eigenvectors of the symbol (5.2), $\zeta A + \omega B$, and then use the ideas of the Proposition to translate these vectors into eigenvectors of (5.5). This would give an exact diagonalization of (5.5) when (ω,ξ) is in Γ, but the expressions for the eigenvectors can be complicated, and in order to obtain local boundary conditions it would usually be necessary to approximate these vectors with simple expressions. We would again use approximations which are valid asymptotically as $\frac{\omega}{\xi} \to 0$. Unlike the perturbation method, which employs one fixed method of approximation, this approach allows the use of various Taylor or rational Padé approximations. Engquist and Majda ([1],[2]) found the latter approximations particularly useful in their work on absorbing boundary conditions for scalar wave equations. The greater flexibility of this approach may be an advantage, in general, but in the example appearing in Part II we are able to use the perturbation method and obtain satisfactory results.

5.4 Summary of the Uncoupling Process

We now outline the uncoupling process in terms of formal manipulations of Fourier transforms. In Part II the method will be translated into the language of pseudo-differential operators and applied in detail to the linearized shallow water equations.

We first solve for w_x in (5.1) to obtain $w_x = A^{-1}w_t - A^{-1}Bw_y - A^{-1}Cw$. In order to simplify the notation we change the meaning of A, B, and C and write the system as

$$w_x = Aw_t + Bw_y + Cw \tag{5.9}$$

The matrix A is diagonal.

When we apply Fourier transforms in y and t to (5.9), the result is

$$\hat{w}_x(x,\omega,\xi) = (i\xi A + i\omega B)\hat{w} + C\hat{w} \tag{5.10}$$

The symbol $\xi A + \omega B$ is the same as the symbol (5.5) discussed earlier, aside from the change in notation. According to the remarks of the preceding sub-section, this symbol is diagonalizable for (ω,ξ) in Γ, and it is only for (ω,ξ) in Γ that we can have rapidly moving waves. For the sake of neatness we use a cutoff function to restrict attention to that set. Let ϕ be a smooth function of ω and ξ which is equal to zero outside Γ and which is equal to 1 on all of Γ except for a thin layer near the boundary. Equation (5.10) can then be written

$$\hat{w}_x(x,\omega,\xi) = (i\xi A + i\omega\phi B)\hat{w} + C\hat{w} + i\omega(1 - \phi)B\hat{w} \tag{5.11}$$

The last term in (5.11) is an error term which we will denote by $E\hat{w}$. It is nonzero only near the edge of Γ, and for fast waves this corresponds to nearly tangential incidence. The error term is therefore insignificant.

The next task is to uncouple the leading order terms. That is, we find a matrix q such that

$$q(i\xi A + i\omega\phi B)q^{-1} \tag{5.12}$$

is closer to diagonal form, or to a suitable block diagonal form, when $\frac{\omega}{\xi}$ is small. This question was discussed in Section 5.3. Denote (5.12) by $g + r$, where g is uncoupled and r is the error in the uncoupling. Equation (5.11) then becomes $\frac{\partial}{\partial x}(q\hat{w}) = q(i\xi A + i\omega\phi B)q^{-1}q\hat{w} +$ $(qC + q_x)\hat{w} + qE\hat{w}$, or

$$\frac{\partial \hat{w}_o}{\partial x} = g\hat{w}_o + (qC + q_x)q^{-1}\hat{w}_o + E_o\hat{w}_o, \tag{5.13}$$

where $\hat{w}_o = q\hat{w}$ and $E_o = r + qEq^{-1}$. $E_o\hat{w}_o$ is an error term which is small when $\frac{\omega}{\xi}$ is close to zero.

Before we proceed further, it is necessary to discuss the degree of homogeneity in ω and ξ, or "order", of various expressions which appear in (5.13). The matrix q can be assumed to be homogeneous of degree zero for large ω and ξ, since we can first uncouple the symbol $i\xi A + i\omega\phi B$ when $\omega^2 + \xi^2 = 1$ and then extend q to be constant along rays from the origin. In the example in Part II we will use one iteration of the perturbation method, so that q will have the form $q = I + \frac{\omega}{\xi}\phi M$. In this case the error in the uncoupling will be $r = O(\xi(\frac{\omega}{\xi}\phi)^2)$. Because q has order zero, (5.12) implies that the uncoupled symbol g must have order one. Similarly, the product $(qC + q_x)q^{-1}$ must have order zero. The system (5.13) is therefore uncoupled to order zero, modulo the insignificant error term $E_o\hat{w}_o$.

For simplicity we will let $z = (qC + q_x)q^{-1}$ and write (5.13) as

$$\frac{\partial \hat{w}_o}{\partial x} = g\hat{w}_o + z\hat{w}_o + E_o\hat{w}_o \tag{5.14}$$

When the coefficients in the system depend on y or t, the coupling of order zero must also include some extra terms generated by the uncoupling of the leading order part of the system. These extra terms will appear in the discussion in Part II.

We now use the method of Section 3.2 to uncouple the terms of order zero. We first multiply (5.14) by a matrix of the form $I + k$, where k is a matrix of order -1 which is to be determined, to obtain

$$\frac{\partial}{\partial x}[(I + k)\hat{w}_o] = (I + k)g(I + k)^{-1}(I + k)\hat{w}_o$$

(5.15)

$$+ (I + k)z\hat{w}_o + (I + k)E_o\hat{w}_o + k_x\hat{w}_o.$$

Let $\hat{w}_1 = (I + k)\hat{w}_o$, $E_1 = (I + k)E_o(I + k)^{-1}$, and write $(I + k)^{-1}$ as $I - k + k^2 - \ldots$. The system (5.15) becomes

$$\frac{\partial \hat{w}_1}{\partial x} = (I + k)g(I - k)\hat{w}_1 + z\hat{w}_1 + \text{order}(-1)\hat{w}_1 + E_1\hat{w}_1$$

$$= g\hat{w}_1 + (kg - gk + z)\hat{w}_1 + \text{order}(-1)\hat{w}_1 + E_1\hat{w}_1$$

To eliminate the coupling of order zero, we choose k so that $kg - gh + z$ is diagonal, or block diagonal. The matrix k will in fact have order -1, as we assumed above, since g has order one and z has order zero. As noted in Section 3.2, this method can be repeated to reduce further the coupling caused by lower order terms.

The new dependent variable is defined by $\hat{w}_1 = (I + k)\hat{w}_o = (I + k)q\hat{w}$. To obtain local boundary conditions, we express the incoming fast components of $\hat{w}_1$ in terms of $\hat{w}$, set these components equal to zero at the boundary, clear denominators, and invert the Fourier transform. An analogue for systems with variable coefficients will be given in Part II.

APPENDIX. A Perturbation Lemma

Here we present a method for reducing the coupling found in matrices which are perturbations of block diagonal matrices. This method can be used to partially uncouple the leading symbol in the system (5.11), and it is essentially the method which has been used to reduce the coupling caused by lower order terms. We present it as a separate lemma for the sake of clarity and generality. Various versions of this method have been used in [8], [11], and [12].

Proposition. Let A and B be square matrices of equal dimension. Suppose that A is block diagonal, and let $A_1,\dots,A_n$ denote the blocks on the diagonal. If no two of the A_j have any eigenvalues in common, then for small ϵ the sum $A + \epsilon B$ can be uncoupled to order ϵ^2. More precisely, there exists a matrix M such that for ϵ sufficiently small,

$$(I + \epsilon M)(A + \epsilon B)(I + \epsilon M)^{-1} = A + \epsilon \cdot (\text{block diagonal matrix}) + O(\epsilon^2).$$

A method for constructing M will be given in the proof.

Proof. For small ϵ the inverse $(I + \epsilon M)^{-1}$ exists and is equal to $I - \epsilon M + \epsilon^2 M^2 - \dots$. We can therefore write

$$\begin{aligned}(I + \epsilon M)(A + \epsilon B)(I + \epsilon M)^{-1} &= (I + \epsilon M)(A + \epsilon B)(I - \epsilon M + O(\epsilon^2)) \\ &= A + \epsilon(MA - AM + B) + O(\epsilon^2).\end{aligned}$$

Our goal is to choose M so that

$$(1) \qquad MA - AM + B$$

is block diagonal. For the sake of notation we will partition M and B

into block structures which match the block structure of A. M_{ij} and B_{ij} will denote the blocks in the (i,j) position. They are not necessarily square, since we are not assuming that A_i and A_j have the same dimensions. The (i,j) block in (1) can then be written as $M_{ij}A_j - A_iM_{ij} + B_{ij}$. For $i \neq j$ we want this to be equal to zero. We are therefore faced with the problem of solving the equation

$$M_{ij}A_j - A_iM_{ij} = -B_{ij} \tag{2}$$

for M_{ij}. Once we have done this, the proof is complete. There are no conditions imposed on the diagonal blocks M_{ii}, so these may be chosen arbitrarily.

If A_i and A_j are both 1×1 matrices, i.e., scalars, then we obviously need to have $A_i \neq A_j$ in order to be able to solve (2) for arbitrary B_{ij}. In the general case the system (2) is solvable if and only if A_i and A_j have disjoint spectra. Proofs of this fact can be found in several different references. We give one here for the sake of completeness.

In order to simplify the notation we will write (2) in the form $XS - TX = Y$, where S, T, and Y are given and X is to be determined. We are assuming that S and T are square matrices which do not have any eigenvalues in common. There is no need to assume that they have the same dimension. We will denote the columns of X and Y by x_j and y_j, and we will denote the entries of S by s_{ij}.

We can assume that S is upper triangular, since otherwise we can use a similarity transformation to reduce the problem to that case. We will solve for the columns of X, starting from the left. We first

have $s_{11}x_1 - Tx_1 = y_1$. The matrix $s_{11}I - T$ is nonsingular since s_{11} is an eigenvalue of S and therefore not an eigenvalue of T. The column x_1 is therefore determined uniquely. We next have $s_{22}x_2 - Tx_2 = y_2 - s_{12}x_1$. This system has a unique solution x_2 since s_{22} is not an eigenvalue of T. We can continue in this manner to solve for all of X. We note that the condition on the eigenvalues of S and T is necessary as well as sufficient, since it is equivalent to the statement that $s_{ii}I - T$ is nonsingular for all i. This completes the proof of the lemma, and therefore the proof of the main proposition.

This method can be applied repeatedly to uncouple the given matrix to any order. To eliminate the coupling of order ϵ^2, we would introduce a matrix of the form $I + \epsilon^2 M_2$ and solve for M_2 in the manner indicated above. In general, to reduce the coupling from order ϵ^n to order ϵ^{n+1}, we would use a transformation of the form $I + \epsilon^n M_n$.

ACKNOWLEDGMENT

This paper is based on material contained in my Ph.D. thesis. I thank my advisor, Joseph Oliger, for the advice and encouragement he has given me.

REFERENCES

[1] B. Engquist and A. Majda, Absorbing boundary conditions for the numerical simulation of waves, Math. Comp., 31(1977), pp. 629-651.

[2] B. Engquist and A. Majda, Radiation boundary conditions for acoustic and elastic wave calculations, Comm. Pure Appl. Math., 32(1979), pp. 313-357.

[3] B. Gustafsson, Asymptotic expansions for hyperbolic problems with different time scales, SIAM J. Numer. Anal., 17(1980), pp. 623-634.

[4] B. Gustafsson, Numerical solution of hyperbolic systems with different time scales using asymptotic expansions, J. Comp. Phys., 36(1980), pp. 209-235.

[5] B. Gustafsson and H.-O. Kreiss, Difference approximations of hyperbolic problems with different time scales I: the reduced problem, Uppsala University, Dept. of Computer Sciences, Report No.86 (1980).

[6] R. Higdon, Boundary Conditions for Hyperbolic Systems of Partial Differential Equations Having Multiple Time Scales, Ph.D. dissertation, Dept. of Mathematics, Stanford University, 1981. Report no. STAN-CS-81-890, Dept. of Comp. Sci., Stanford, 1981.

[7] R. Higdon, Boundary conditions for suppressing rapidly moving components in hyperbolic systems II: an example, M.R.C. Technical Summary Report #2404.

[8] H.-O. Kreiss, Problems with different time scales for ordinary differential equations, SIAM J. Numer. Anal., 16(1979), pp. 980-998.

[9] H.-O. Kreiss, Problems with different time scales for partial differential equations, Comm. Pure Appl. Math., 33(1980), pp. 399-439.

[10] L. Nirenberg, Lectures on Linear Partial Differential Equations, C.B.M.S. Regional Conf. Ser. in Math., No. 17, Amer. Math. Soc., Providence, RI, 1973.

[11] R.E. O'Malley, Jr., and L.R. Anderson, Singular perturbations, order reduction, and decoupling of large scale systems, Numerical Analysis of Singular Perturbation Problems, P. Hemker and J. Miller, eds., Academic Press, London, 1979, pp. 317-338.

[12] M. Taylor, Reflections of singularities of solutions to systems of differential equations, Comm. Pure Appl. Math., 28(1975), pp. 457-478.

RLH/jk

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: 2403	2. GOVT ACCESSION NO.: AD-A120 962	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Boundary Conditions for Suppressing Rapidly Moving Components in Hyperbolic Systems, I		5. TYPE OF REPORT & PERIOD COVERED: Summary Report - no specific reporting period 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Robert L. Higdon		8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-1132, DAAG29-80-C-0041, MCS-7927062, Mod. 1
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Work Unit Number 3 - Numerical Analysis and Computer Science
11. CONTROLLING OFFICE NAME AND ADDRESS: See Item 18 below		12. REPORT DATE: July 1982 13. NUMBER OF PAGES: 40
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report): UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U.S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709
National Science Foundation, Washington, D.C. 20550
Office of Naval Research, Arlington, VA 22217

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

hyperbolic systems, absorbing boundary conditions, multiple time scales, reflection of singularities.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This work is concerned with hyperbolic systems of partial differential equations for which certain of the associated propagation speeds are a great deal larger than the other propagation speeds. Our goal is to find boundary conditions which prevent rapidly moving waves from entering the given spatial domain. Conditions of this type are desirable in certain numerical computations arising in meteorology.

In order to find these conditions, we first transform the given system to an approximate diagonal form in such a way that each of the new dependent

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT (cont.)

variables can be identified as a slow, incoming fast, or outgoing fast component of the solution. We then find local boundary conditions which suppress the incoming fast part. Pseudo-differential operators are used throughout the entire process. We consider only linear systems. In Part II of this work (7) these methods are applied in detail to the linearized shallow water equations.

The results of numerical tests of various boundary conditions are included in both papers. We also outline a method by which the conditions can be justified analytically.